# Moyennes, itérations et convergence quadratique

**Stéphane Junca**

*Laboratoire J. A. Dieudonné, UMR CNRS 6621,*
*Université de Nice-Sophia Antipolis,*
*Parc Valrose, B.P. 71, F-06108 Nice, Cédex 2, France,*

On étend à des moyennes générales ce qui est connu pour la moyenne arithmético-géométrique, avec deux variables puis avec un nombre quelconque de variables.

En général, il est très classique de montrer qu'une suite itérative $X_{n+1} = F(X_n)$ converge très rapidement vers un point fixe $L$ de $F$ si la différentielle de $F$ en $L$ est nulle : $DF(L) = 0$. Ici, on donne une famille de fonctions $F$ pour laquelle la différentielle évaluée en $L$ n'est pas nulle mais dont les suites itératives convergent très rapidement. En fait, l'application $F$ n'est même pas contractante (le rayon spectral de $DF(L)$ étant égal à 1). Et cependant la direction associée à la valeur propre nulle de $DF(L)$ est super attractive.

Les suites arithmético-géométrique, arithmético-harmonique et harmonico-géométrique sont des exemples bien connus de ce phénomène étrange de super-convergence.

## 1 Dans $\mathbb{R}^2$

**Définition 1 (Moyenne)**
*Soit $I = ]0, +\infty[$ $\mu$ est une moyenne si $\mu$ est une application de $I \times I$ dans $I$ telle que :*

- *(i)* $\forall (x,y) \in I \times I,\ \mu(x,y) = \mu(y,x)$
- *(ii)* $\forall (x,y) \in I \times I,\ \min(x,y) \leq \mu(x,y) \leq \max(x,y)$

*On dit que la moyenne est stricte si*

- *(iii)* $\forall (x,y) \in I \times I,\ \mu(x,y) = x \Longleftrightarrow x = y.$

Exemples de moyennes strictes : $\sqrt{xy}$, $\left(\dfrac{x^p + y^p}{2}\right)^{\frac{1}{p}}$ où $p \neq 0$, voir [1], ...

**Définition 2 ( Convergence quadratique)**
*Soit $d \in \mathbb{N} - \{0\}$, $(X_n)_{n \in \mathbb{N}}$ une suite de $\mathbb{R}^d$ convergeant vers $L$. On dit que la convergence est quadratique si $X_{n+1} - L = O(\|X_n - L\|^2)$ pour tout $n \in \mathbb{N}$, où $\|.\|$ est une norme de $\mathbb{R}^d$.*

La convergence quadratique est très rapide. En pratique, elle double le nombre de décimales à chaque itération. Ainsi, elle sature la précision d'une calculette en 3 ou 4 itérations. La suite de Héron définie par $x_{n+1} = \dfrac{1}{2}\left(x_n + \dfrac{2}{x_n}\right)$ pour $n \in \mathbb{N}$ et $x_0 = 1$, a une convergence quadratique vers $\sqrt{2}$. Plus généralement, la méthode de Newton permet d'approcher la racine d'une équation avec une convergence quadratique.

**Théorème 1**
*Soient $\mu$ et $\nu$ deux moyennes de classe $C^2$ telles que : l'une des deux moyennes est stricte et*

$$\forall (x,y) \in I \times I,\ \mu(x,y) \geq \nu(x,y). \tag{1}$$

*Soit* $F(x,y) = \begin{pmatrix} \mu(x,y) \\ \nu(x,y) \end{pmatrix}$ *et* $X_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$ *la suite définie par*

$$\left\{ \begin{array}{lcll} X_{n+1} & = & F(X_n) & n \in \mathbb{N} \\ X_0 & = & \begin{pmatrix} a \\ b \end{pmatrix} & \in I \times I \end{array} \right.$$

*Alors les suites $(x_n)_{n\geq 1}$ et $(y_n)_{n\geq 1}$ sont adjacentes. Leur limite $m(a,b)$ est une moyenne. Si $\mu$ et $\nu$ sont strictes alors $m$ est stricte.* La convergence est quadratique.

Les propriétés de $m(.,.)$ sont évidentes. En revanche, la convergence quadratique des suites est plutôt surprenante.

**Remarque 1** *Si $\mu$ $\nu$ ne sont pas strictes, $(x_n)_{n\geq 1}$ et $(y_n)_{n\geq 1}$ ne sont plus nécessairement adjacentes. En revanche la convergence reste quadratique.*

Démonstration : D'après (1) $x_1 \geq y_1$, puis, par récurrence sur $n$ d'après (1) et (ii) on obtient pour $n \geq 1$, $\max(a,b) \geq x_1 \geq x_n \geq x_{n+1} \geq y_{n+1} \geq y_n \geq y_1 \geq \min(a,b)$. Donc $(x_n)_{n\geq 1}$ et $(y_n)_{n\geq 1}$ sont des suites monotones et convergent respectivement vers $x$ et $y$ appartenant à $[\min(a,b), \max(a,b)]$. Par continuité de $F$, on a $x = \mu(x,y)$ et $y = \nu(x,y)$. Comme l'une des deux moyennes est stricte on a forcément $x = y$. Donc ces suites sont adjacentes.

Démontrons, maintenant, que la convergence est quadratique. Comme $\mu(x,x) = x$ on a en dérivant et en utilisant la symétrie $\dfrac{\partial \mu}{\partial x}(x,x) + \dfrac{\partial \mu}{\partial y}(x,x) = 1$ et $\dfrac{\partial \mu}{\partial x}(x,x) = \dfrac{\partial \mu}{\partial y}(x,x)$. Ainsi la matrice de la différentielle de $F$ sur la première bissectrice est

$$A = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Notons $e_n = x_n - m(a,b)$, $f_n = y_n - m(a,b)$ et $L = m(a,b)\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Le développement de Taylor de $F$ en $L$ donne $X_{n+1} - L = F(X_n) - F(L) = A(X_n - L) + O(e_n^2 + f_n^2)$. En utilisant le fait que : $e_{n+1} \geq 0 \geq f_{n+1}$ car $(x_n)_{n\geq 1}$ et $(y_n)_{n\geq 1}$ sont adjacentes, on a

$$0 \leq e_{n+1} = \frac{1}{2}(e_n + f_n) + O(e_n^2 + f_n^2) \tag{2}$$

$$0 \geq f_{n+1} = \frac{1}{2}(e_n + f_n) + O(e_n^2 + f_n^2) \tag{3}$$

On en déduit l'encadrement suivant : $\frac{1}{2}(e_n + f_n) = O(e_n^2 + f_n^2)$. En reportant cette dernière estimation dans (2) et (3) on obtient $e_{n+1} = O(e_n^2 + f_n^2)$ et $f_{n+1} = O(e_n^2 + f_n^2)$. □

**Remarque 2**
*Notons $E_n = X_n - L$, $Q = \{(x,y) \in \mathbb{R} \times \mathbb{R} \ / \ x > 0$ et $y < 0\}$ et $D = \{(x,x) \in \mathbb{R} \times \mathbb{R}\}$. Pour tout $n \geq 1$, $E_n$ appartient à $Q$. D'autre part, $A$ est la matrice de la projection orthogonale sur $D$. Ainsi $AE_n$ appartient à $D$. Mais $D \cap Q = \emptyset$ et pour $Y \in D$ la distance euclidienne de $Y$ à $Q$ est égale à la norme euclidienne de $Y$ divisée par $\sqrt{2}$. Or $E_{n+1} = AE_n + O(\|E_n\|^2)$ appartient à $Q$. On a donc forcément que la distance de $AE_n$ à $Q$ est en $O(\|E_n\|^2)$. Ce qui nous donne $AE_n = O(\|E_n\|^2)$ et ainsi $E_{n+1} = O(\|E_n\|^2)$.*

## 2 Généralisations dans $\mathbb{R}^d$

On suppose que $d \geq 3$. On note $S_d$ le groupe des permutations de $\{1, \cdots, d\}$.

**Définition 3 ( moyenne à $d$ variables)**
*$\mu$ est une moyenne à $d$ variables si $\mu$ est une application de $I^d$ dans $I$ telle que :*

- *(i)* $\forall x = (x_1, \cdots, x_d) \in I^d, \ \forall \sigma \in S_d, \ \mu(x_{\sigma(1)}, \cdots, x_{\sigma(d)}) = \mu(x_1, \cdots, x_d)$
- *(ii)* $\forall x \in I^d, \ \min_{i=1}^{d} x_i \leq \mu(x) \leq \max_{i=1}^{d} x_i$

*On dit que la moyenne est stricte si*

- *(iv)* $\forall x = (x_1, \cdots, x_d) \in I^d, \ \mu(x) = x_1 \Longleftrightarrow x_1 = \cdots = x_d$.

Exemples de moyennes à 3 variables strictes : $(x_1 x_2 x_3)^{\frac{1}{3}}$, $\left(\dfrac{x_1^p + x_2^p + x_3^p}{3}\right)^{\frac{1}{p}}$ où $p \neq 0, \ldots$

**Théorème 2**
*Soit $\mu_1, \cdots, \mu_d$ des moyennes à $d$ variables de classe $C^2$ telles que : $\mu_1$ ou $\mu_d$ est stricte et*

$$\forall k \in \{2, \cdots, d-1\}, \ \forall x \in I^d, \ \mu_1(x) \geq \mu_k(x) \geq \mu_d(x). \tag{4}$$

*Soit* $F(x) = \begin{pmatrix} \mu_1(x) \\ \vdots \\ \mu_d(x) \end{pmatrix}$ *et* $X_n = \begin{pmatrix} x_{1,n} \\ \vdots \\ x_{d,n} \end{pmatrix}$ *la suite définie par*

$$\begin{cases} X_{n+1} & = & F(X_n) & n \in \mathbb{N} \\ X_0 & = & \begin{pmatrix} a_1 \\ \vdots \\ a_d \end{pmatrix} & \in I^d \end{cases}$$

*Alors les suites $(x_{1,n})_{n \geq 1}$ et $(x_{d,n})_{n \geq 1}$ sont adjacentes de limite $m(a)$. $m(.)$ est une moyenne. $(X_n)_{n \in \mathbb{N}}$ converge vers un élément de la diagonale : $D^+ = \{x \in I^d \ / x_1 = \cdots = x_d\}$ et* la convergence est quadratique.

La moyenne arithmético-géométrico-harmonique est un exemple classique de convergence quadratique.

DÉMONSTRATION : Supposons que $\mu_1$ est stricte. Notons $\alpha = \max_{i=1}^{d} a_i$ et $\beta = \min_{i=1}^{d} a_i$. On montre, comme dans la démonstration du théorème 1, que

$$\forall n \geq 1,\ \forall k \in \{2, \cdots, d-1\}, \quad \alpha \geq x_{1,1} \geq x_{1,n} \geq x_{1,n+1} \geq x_{1,n+1} \geq x_{d,n} \geq x_{d,1} \geq \beta, \tag{5}$$
$$\forall n \geq 1,\ \forall k \in \{2, \cdots, d-1\}, \quad x_{1,n} \geq x_{k,n} \geq x_{d,n} \tag{6}$$

Donc $(x_{1,n})_{n\geq 1}$ et $(x_{d,n})_{n\geq 1}$ sont des suites monotones et convergent respectivement vers $x_1$ et $x_d$ appartenant à $I$. Pour tout $k \in \{2, \cdots, d-1\}$ la suite $(x_{kn)})_{n \geq 1}$ est contenue dans $[\beta, \alpha]$ d'après (6). Donc on peut extraire une sous-suite telle que, pour tout $k \in \{2, \cdots, d-1\}$, $(x_{k\phi(n)})_{n\geq 1}$ converge vers $x_k \in [\beta, \alpha]$. Remarquons que $(x_{1,\phi(n)+1})_{n\geq 1}$ converge aussi vers $x_1$. Donc, par continuité de $\mu_1$ on a $x_1 = \mu_1(x_1, \cdots, x_d)$. Comme $\mu_1$ est stricte on a forcément $x_1 = \cdots = x_d$. Donc,$(x_{1,n})_{n\geq 1}$ et $(x_{d,n})_{n\geq 1}$ sont des suites adjacentes. Et, par (6), $(X_n)$ converge vers $L \in D^+$.

Il reste à montrer que la convergence est quadratique. Par symétrie et du fait que $F$ est l'identité sur $D^+$, la différentielle de $F$ en $L$ est :

$$A = \frac{1}{d} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}.$$

Notons $e_{in} = x_{in} - l$, où $l$ est une composante de $L$, et $\varepsilon_n = \left( \sum_{i=1}^{d} e_{in}^2 \right)^{\frac{1}{2}}$. Le développement de Taylor de $F$ en $L$ donne $X_{n+1} - L = F(X_n) - F(L) = A(X_n - L) + O(\varepsilon_n^2)$. On a pour $i = 1, \ldots, d$

$$e_{i,n+1} \quad = \quad \frac{1}{d} \sum_{j=1}^{d} e_{j,n} + O(\varepsilon_n^2) \tag{7}$$

De $e_{1,n+1} \geq 0 \geq e_{d,n+1}$ on tire, $\frac{1}{d} \sum_{j=1}^{d} e_{j,n} = O(\varepsilon_n^2)$, donc, en revenant à (7), $\varepsilon_{n+1} = O(\varepsilon_n^2)$.

□

**Remarque 3**
*On a une interprétation géométrique similaire au cas où $d = 2$. Notons $E_n = X_n - L$, $Q = \{x \in \mathbb{R}^d \ / \ x_1 > 0$ et $x_d < 0\}$ et $D$ la droite contenant $D^+$. Pour tout $n \geq 1$, $E_n$ appartient à $Q$. Comme $A$ est la matrice de la projection orthogonale sur $D$, et que la distance euclidienne d'un élément de $D$ à $Q$ est proportionnelle à la norme euclidienne de cet élément, on a nécessairement $AE_n = O(\|E_n\|^2)$. Ce qui nous donne le résultat.*

## Références


[1] Gustave Choquet, Cours de Topologie, Masson ( 2ème éd ), (1992), p. 161-167.

[2] Jean-Pierre Demailly, Analyse numérique et équations différentielles, Presse Universitaire de Grenoble, (1991), p. 95-115.



Je remercie François Rouvière pour l'intérêt qu'il a manifesté pour ce papier.